

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re the Application of:

EUNG DON LEE, ET AL.

Application No.:

Filed:

For: **APPARATUS FOR AND METHOD OF
DETERMINING TRANSMISSION RATE
IN SPEECH TRANSCODING**

Art Group:

Examiner:

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

REQUEST FOR PRIORITY

Sir:

Applicant respectfully requests a convention priority for the above-captioned application, namely:

<u>COUNTRY</u>	<u>APPLICATION NUMBER</u>	<u>DATE OF FILING</u>
Korea	10-2003-0043374	30 June 2003

☒ A certified copy of the document is being submitted herewith.

Respectfully submitted,

Blakely, Sokoloff, Taylor & Zafman LLP

Dated: 12/4/03

12400 Wilshire Boulevard, 7th Floor
Los Angeles, CA 90025
Telephone: (310) 207-3800


Eric S. Hyman, Reg. No. 30,139

**KOREAN INTELLECTUAL
PROPERTY OFFICE**

This is to certify that the following application annexed hereto is a true copy from the records of the Korean Intellectual Property Office.

Application Number: Korean Patent Application 2003-0043374

Date of Application: 30 June 2003

Applicant(s): Electronics and Telecommunications Research Institute

21 November 2003

COMMISSIONER

[Bibliography]

[Document Name]	Patent Application
[Classification]	Patent
[Receiver]	Commissioner
[Reference No.]	0001
[Filing Date]	30 June 2003
[IPC]	H03M
[Title]	Apparatus and method for determining transmission rate in speech code transcoding
[Applicant]	
[Name]	Electronics and Telecommunications Research Institute
[Applicant code]	3-1998-007763-8
[Attorney]	
[Name]	Youngpil Lee
[Attorney code]	9-1998-000334-6
[General Power of Attorney Registration No.]	2001-038378-6
[Attorney]	
[Name]	Haeyoung Lee
[Attorney code]	9-1999-000227-4
[General Power of Attorney Registration No.]	2001-038396-8
[Inventor]	
[Name]	LEE, Eung Don
[Resident Registration No.]	710713-1684311
[Zip Code]	305-390
[Address]	304-902 Expo Apt., Jeonmin-dong, Yusong-gu Daejeon-city, Rep. of Korea
[Nationality]	Republic of Korea
[Inventor]	
[Name]	KIM, Hyun Woo
[Resident Registration No.]	781121-1047124
[Zip Code]	143-773
[Address]	701-1608 Woosung Apt., Jayang-dong, Gwangjin-gu Seoul, Rep. of Korea
[Nationality]	Republic of Korea
[Inventor]	
[Name]	KIM, Do Young
[Resident Registration No.]	600830-1047810

[Zip Code] 305-755
[Address] 118-1404 Hanbit Apt., Eoeun-duong, Yusong-gu
Daejeon-city, Rep. of Korea
[Nationality] Republic of Korea

[Inventor]
[Name] YOO, Chang Dong
[Resident]
Registration No.] 631109-1002511
[Zip Code] 305-340
[Address] 1-503 KAIST Prof. Apt., Doryong-dong, Yusong-gu
Daejeon-city, Rep. of Korea
[Nationality] Republic of Korea

[Inventor]
[Name] SEO, Seong Ho
[Resident]
Registration No.] 790620-1852739
[Zip Code] 711-813
[Address] 101-1502 Bosung Town, Seojae-ri, Dasa-eup, Dalsung-gun
Daegu-city, Rep. of Korea
[Nationality] Republic of Korea

[Inventor]
[Name] JANG, Dal Won
[Resident]
Registration No.] 800326-1840810
[Zip Code] 645-230
[Address] 2-11 Chungmu-dong, Jinhae-city
Kyungsangnam-do, Rep. of Korea
[Nationality] Republic of Korea

[Request for
Examination] Requested

[Purpose] We file as above according to Art. 42 of the Patent Law,
request the examination as above according to Art. 60 of the
Patent Law.

Attorney Youngpil Lee
Attorney Haeyoung Lee

[Fee]
[Basic page] 20 Sheet(s) 29,000 won
[Additional page] 1 Sheet(S) 1,000 won
[Priority claiming fee] 0 Case(S) 0 won
[Examination fee] 9 Claim(s) 397,000 won
[Total] 427,000 won

1020030043374

Print Date: 2003/11/27

[Reason for Reduction]	Government Invented Research Institution
[Fee after Reduction]	213,500 won

[Transfer of Technology]	Allowable
[Licensing]	Allowable
[Technology Training]	Allowable

[Enclosures]

1. Abstract and Specification (and Drawings)	1 copy
----------------------------------------------	--------



별첨 사본은 아래 출원의 원본과 동일함을 증명함.

This is to certify that the following application annexed hereto is a true copy from the records of the Korean Intellectual Property Office.

출원 번호 : 10-2003-0043374
Application Number

출원 년 월 일 : 2003년 06월 30일
Date of Application JUN 30, 2003

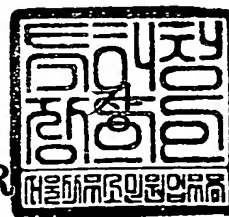
출원인 : 한국전자통신연구원
Applicant(s) Electronics and Telecommunications Research Insti



2003 년 11 월 21 일

특 허 청

COMMISSIONER



【서지사항】

【서류명】	특허출원서
【권리구분】	특허
【수신처】	특허청장
【참조번호】	0001
【제출일자】	2003.06.30
【국제특허분류】	H03M
【발명의 명칭】	음성의 상호부호화시 전송률 결정 장치 및 방법
【발명의 영문명칭】	Apparatus and method for determining transmission rate in speech code transcoding
【출원인】	
【명칭】	한국전자통신연구원
【출원인코드】	3-1998-007763-8
【대리인】	
【성명】	이영필
【대리인코드】	9-1998-000334-6
【포괄위임등록번호】	2001-038378-6
【대리인】	
【성명】	이해영
【대리인코드】	9-1999-000227-4
【포괄위임등록번호】	2001-038396-8
【발명자】	
【성명의 국문표기】	이응돈
【성명의 영문표기】	LEE,Eung Don
【주민등록번호】	710713-1684311
【우편번호】	305-390
【주소】	대전광역시 유성구 전민동 엑스포아파트 304동 902호
【국적】	KR
【발명자】	
【성명의 국문표기】	김현우
【성명의 영문표기】	KIM,Hyun Woo
【주민등록번호】	781121-1047124

【우편번호】	143-773
【주소】	서울특별시 광진구 자양3동 우성7차아파트 701동 1608호
【국적】	KR
【발명자】	
【성명의 국문표기】	김도영
【성명의 영문표기】	KIM,Do Young
【주민등록번호】	600830-1047810
【우편번호】	305-755
【주소】	대전광역시 유성구 어은동 한빛아파트 118동 1404호
【국적】	KR
【발명자】	
【성명의 국문표기】	유창동
【성명의 영문표기】	Y00,Chang Dong
【주민등록번호】	631109-1002511
【우편번호】	305-340
【주소】	대전광역시 유성구 도룡동 과기원 교수아파트 1동 503호
【국적】	KR
【발명자】	
【성명의 국문표기】	서성호
【성명의 영문표기】	SE0,Seong Ho
【주민등록번호】	790620-1852739
【우편번호】	711-813
【주소】	대구광역시 달성군 다사읍 서재리 145 서재보성타운 101동 1502호
【국적】	KR
【발명자】	
【성명의 국문표기】	장달원
【성명의 영문표기】	JANG,Dal Won
【주민등록번호】	800326-1840810
【우편번호】	645-230
【주소】	경상남도 진해시 충무동 2-11
【국적】	KR
【심사청구】	청구

【취지】

특허법 제42조의 규정에 의한 출원, 특허법 제60조의 규정에 의한 출원심사를 청구합니다. 대리인

이영필 (인) 대리인

이해영 (인)

【수수료】

【기본출원료】 20 면 29,000 원

【가산출원료】 1 면 1,000 원

【우선권주장료】 0 건 0 원

【심사청구료】 9 항 397,000 원

【합계】 427,000 원

【감면사유】 정부출연연구기관

【감면후 수수료】 213,500 원

【기술이전】

【기술양도】 희망

【실시권 허여】 희망

【기술지도】 희망

【첨부서류】

1. 요약서·명세서(도면)_1통

【요약서】

【요약】

음성의 상호부호화시 전송률 결정 장치 및 방법이 개시된다. 음성/묵음분류부는 부호화된 비트열의 입력 파라미터로부터 소정의 제1문턱값을 기초로 입력 프레임을 음성과 묵음으로 분류한다. 유/무성음분류부는 입력 파라미터에 대응하는 적응 코드북 이득에 대해 설정된 제2문턱값을 기초로 음성으로 분류된 입력 파라미터로부터 입력 프레임을 유성음과 무성음으로 분류한다. 유성음/변화음분류부는 과거 프레임의 클래스를 기초로 유/무성음분류부에 의해 유성음으로 분류된 입력 파라미터로부터 입력 프레임을 유성음과 변화음으로 분류한다. 유성음분류부는 적응 코드북 이득의 변화량 또는 입력 프레임에 대한 피치지연값의 최대값과 최소값의 차이값에 대해 설정된 소정의 제3문턱값을 기초로 유성음/변화음분류부에 의해 유성음으로 분류된 입력 파라미터로부터 입력 프레임을 일정한 유성음과 일정하지 않은 유성음으로 분류한다. 전송률결정부는 분류결과에 대응하는 클래스에 대해 설정되어 있는 전송률과 형태를 기초로 입력 프레임에 대한 전송률을 결정한다. 본 발명에 따르면, 프레임을 용이하게 분류하고 전송률 결정 과정을 간단하게 구현할 수 있으며, 전체적인 계산량이 감소된다.

【대표도】

도 2

【색인어】

CELP, SMV, 음성부호화, 상호부호화, Transcoding

【명세서】

【발명의 명칭】

음성의 상호부호화시 전송률 결정 장치 및 방법{Apparatus and method for determining transmission rate in speech code transcoding}

【도면의 간단한 설명】

도 1은 종래의 SMV 음성부호화기에서의 전송률 결정과정을 도시한 도면,

도 2는 본 발명에 따른 음성의 상호부호화시 전송률 결정 장치의 구성을 도시한 블록도,

도 3은 일정 구간 동안(두 프레임) 입력으로 들어오는 G.729A 신호의 피치지연값의 최대값과 최소값의 차이 및 사용한 음성을 도시한 도면,

도 4는 적응 코드북 이득값의 프레임별 최소값을 도시한 도면,

도 5는 단일음성신호에 대해서 깨끗한 신호와 백색잡음을 섞은 신호에 대해서 G.729A의 고정 코드북 이득값을 도시한 도면, 그리고,

도 6은 본 발명에 따른 음성의 상호부호화시 전송률 결정 방법의 수행과정을 도시한 흐름도이다.

【발명의 상세한 설명】

【발명의 목적】

【발명이 속하는 기술분야 및 그 분야의 종래기술】

<7> 본 발명은 음성 상호부호화시 전송률 결정장치 및 방법에 관한 것으로, 보다 상세하게는, CELP 기반의 음성부호화기로 부호화된 신호를 SMV 신호로 상호부호화할 때 전송률을 결정하는 장치 및 방법에 관한 것이다.

<8> 음성 상호부호화는 하나의 부호화기에서 부호화된 비트열을 또 다른 음성부호화의 비트열로 변환하는 과정을 말한다. 음성 상호부호화 장치는 하나의 음성코덱의 복호화기(decoder)와 다른 음성코덱의 부호화기(encoder)를 그대로 연결함으로써 구현될 수 있다. 그러나, 복호화기와 부호화기를 직접 연결하는 방식은 상호부호화로 인한 지연 시간이 커지고, 계산량이 많아지는 문제점이 있다. 이러한 문제를 해결하기 위해서 음성을 완전히 복호화하지 않고 파라미터 단계에서 직접 변환을 하는 상호부호화기가 각 음성부호화기 쌍에 대해서 개발되고 있다.

<9> 현재, 여러 가지 음성부호화기가 다양한 통신환경에 맞추어 표준화되어서 사용되고 있다. CDMA(Code Division Multiple Access)방식에서 표준 음성부호화기로는 SMV(Selected Mode Vocoder)가 사용된다. SMV는 대역폭(bandwidth)을 절약하기 위해서 각 프레임마다 전송률을 결정한다. SMV 음성부호화기는 8.55, 4.0, 2.0, 및 0.8kbps의 전송률을 가지고 있으며, 프레임 단위로 전송률을 결정해서 부호화한다. 위의 네 개의 전송률을 각각 Rate 1(full-rate), Rate 1/2(half-rate), Rate 1/4(quarter-rate), Rate 1/8(eighth-rate)라고 한다. 그리고, Rate 1과 Rate 1/2은 type 0과 type 1의 두가지 형태를 취할 수 있다. 프레임이 일정한 유성음 구간이면 type 1에 해당되고, 그 외의 경위는 type 0에 해당된다. 전송률과 타입을 결정하기 위해서 SMV 음성부호화기는 입력을 총 6가지의 프레임 클래스(frame class)로 분류한다. 이 과정을 프레임 분류(frame classification) 과정이라고 한다. 6가지 프레임 클래스는 묵음(silence), 잡음(noise-like), 무성음(unvoiced), 변화음(onset), 일정하지 않은 유성음(non-stationary voiced), 일정한 유성음(stationary voiced)이다.

<10> 도 1은 종래의 SMV 음성부호화기에서의 전송률 결정과정을 도시한 도면이다.

<11> 도 1을 참조하면, SMV 음성부호화기로 입력된 음성신호에 대해 전처리과정이 수행된다(S100). 전처리된 음성신호로부터 선형예측계수(Linear Prediction Coefficient : LPC)가 산출

되며(S110), 전처리된 음성신호와 선형예측계수에 대해 인식가중필터링(Perceptual Weighting Filter)이 수행된다(S120). 한편, 선형예측계수로부터 음성영역검출(Voice Activity Detection)이 수행되고(S130), 선형예측계수와 검출된 음성영역으로부터 음악검출이 이루어진다(S140). 또한, 인식가중필터링이 수행된 선형예측계수에 대해 유무성음 레벨이 결정되고(S150), 선형예측계수와 인식가중필터링이 수행된 선형예측계수로부터 개방루프 피치 검출이 수행된다(S160). 마지막으로 검출된 개방루프 피치, 유무성음 레벨, 음악검출결과, 음성영역검출결과, 및 선형예측계수를 설정되어 있는 문턱값(threshold value)과 비교하여 프레임이 속하는 클래스가 결정되고 결정된 클래스에 해당하는 전송률이 결정된다(S170). 표 1에는 프레임이 속하는 클래스에 해당하는 전송률이 기재되어 있다.

<12> 【표 1】

모드	프레임 클래스	레이트 $\frac{1}{8}$	레이트 $\frac{1}{4}$	레이트 $\frac{1}{2}$	레이트 1
0	목음	√			
	잡음			√	√
	무성음			√	√
	변화음			√	√
	일정하지 않은 유성음				√
	일정한 유성음				√
1,2,3	목음	√			
	잡음		√	√	
	무성음		√	√	
	변화음		√	√	√
	일정하지 않은 유성음			√	√
	일정한 유성음			√	√

<13> 상술한 SMV 음성부호화기의 전송률 결정 과정을 상호부호화기에서 사용할 경우 다음과 같은 문제점이 존재한다.

<14> 첫째, SMV 음성부호화기의 전송률 결정 알고리즘은 입력된 음성으로부터 계산된 여러가지 음성 파라미터를 기초로 전송률을 결정한다. 그러나, 상호부호화기로 입력되는 신호는 음성이 아닌 비트열이라는 문제가 있다.

<15> 둘째, 도 1에 도시되어 있는 바와 같이 전송률 결정 과정에서 선형예측분석 과정(LP analysis)이나 개방루프 피치검색과정(open loop pitch detection)이 필요하지만 상호부호화기에서는 이러한 과정들이 필요하지 않다는 문제가 있다. 따라서, 상호부호화기에 SMV의 전송률 결정 과정을 적용하는 것은 가능하나 효율성이 떨어진다.

【발명이 이루고자 하는 기술적 과제】

<16> 본 발명이 이루고자 하는 기술적 과제는 CELP 기반의 다른 음성부호화기로 부호화된 신호를 SMV 음성부호화기의 신호로 바꾸는 상호부호화기에서 입력되는 비트열의 파라미터를 이용하여 전송률을 결정할 수 있는 장치 및 방법을 제공하는 데 있다.

<17> 본 발명이 이루고자 하는 기술적 과제는 CELP 기반의 다른 음성부호화기로 부호화된 신호를 SMV 음성부호화기의 신호로 바꾸는 상호부호화기에서 입력되는 비트열의 파라미터를 이용하여 전송률을 결정할 수 있는 방법을 컴퓨터에서 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체를 제공하는 데 있다.

【발명의 구성 및 작용】

<18> 상기의 기술적 과제를 달성하기 위한, 본 발명에 따른 음성의 상호부호화시 전송률 결정 장치는, 부호화된 비트열의 입력 파라미터에 대응하는 고정 코드북 이득, 적응 코드북 이득, 잡음대 신호비, 및 피치지연 중에서 적어도 하나의 값에 대해 설정된 제1문턱값을 기초로 상기 입력 파라미터로부터 입력 프레임을 음성과 묵음으로 분류하는 음성/묵음분류부; 상기 적응 코

드북 이득에 대해 설정된 제2문턱값을 기초로 음성으로 분류된 상기 입력 파라미터로부터 상기 입력 프레임의 유성음과 무성음으로 분류하는 유/무성음분류부; 과거 프레임의 클래스를 기초로 상기 유/무성음분류부에 의해 유성음으로 분류된 상기 입력 파라미터로부터 상기 입력 프레임을 유성음과 변화음으로 분류하는 유성음/변화음분류부; 상기 적응 코드북 이득의 변화량 또는 상기 피치지연의 최대값과 최소값의 차이에 대해 설정된 제3문턱값을 기초로 상기 유성음/변화음분류부에 의해 유성음으로 분류된 상기 입력 파라미터로부터 상기 입력 프레임을 일정한 유성음과 일정하지 않은 유성음으로 분류하는 유성음분류부; 및 상기 분류결과에 대응하는 클래스에 대해 설정되어 있는 전송률과 형태를 기초로 상기 입력 프레임에 대한 전송률을 결정하는 전송률결정부;를 갖는다.

<19> 상기의 기술적 과제를 달성하기 위한, 본 발명에 따른 음성의 상호부호화시 전송률 결정 방법은, (a) 부호화된 비트열의 입력 파라미터로부터 소정의 제1문턱값을 기초로 입력 프레임을 음성과 묵음으로 분류하는 단계; (b) 상기 입력 파라미터에 대응하는 적응 코드북 이득에 대해 설정된 제2문턱값을 기초로 음성으로 분류된 상기 입력 파라미터로부터 상기 입력 프레임을 유성음과 무성음으로 분류하는 단계; (c) 과거 프레임의 클래스를 기초로 상기 유/무성음분류부에 의해 유성음으로 분류된 상기 입력 파라미터로부터 상기 입력 프레임을 유성음과 변화음으로 분류하는 단계; (d) 상기 적응 코드북 이득의 변화량 또는 상기 입력 프레임에 대한 상기 피치지연값의 최대값과 최소값의 차이값에 대해 설정된 소정의 제3문턱값을 기초로 상기 유성음/변화음분류부에 의해 유성음으로 분류된 상기 입력 파라미터로부터 상기 입력 프레임을 일정한 유성음과 일정하지 않은 유성음으로 분류하는 단계; 및 (e) 상기 분류결과에 대응하는 클래스에 대해 설정되어 있는 전송률과 형태를 기초로 상기 입력 프레임에 대한 전송률을 결정하는 단계;를 포함한다.

- <20> 이에 의해, 프레임을 용이하게 분류하고 전송률 결정 과정을 간단하게 구현할 수 있으며, 전체적인 계산량이 감소될 수 있다.
- <21> 이하에서, 첨부된 도면들을 참조하여 본 발명에 따른 음성의 상호부호화시 전송률 결정 장치 및 방법의 바람직한 실시예를 상세하게 설명한다.
- <22> 도 2는 본 발명에 따른 음성의 상호부호화시 전송률 결정 장치의 구성을 도시한 블록도이다. 종래의 SMV 음성부호화기에서는 전송률 결정을 위해서 총 6개의 종류로 프레임을 분류한다. 본 발명에 따른 음성의 상호부호화시 전송률 결정 장치는 간단함을 위해서 잡음과 무성음을 무성음으로 합쳐서 5가지로 분류한다. 또한, 도 2에 도시된 음성의 상호부호화시 전송률 결정 장치는 G.729A에서 SMV로 상호부호화하는 경우에 전송률을 결정하는 경우를 예로 든 것이다. 다른 코덱에 대해서는 분류 기준이 달라질 수 있으며, 이하에서는 G.729A를 SMV로 상호부호화하는 경우를 예로 들어 설명한다.
- <23> 도 2를 참조하면, 본 발명에 따른 음성의 상호부호화시 전송률 결정 장치는, 음성/목음 분류부(210), 유/무성음분류부(220), 유성음/변화음분류부(230), 유성음분류부(240), 및 전송률결정부(250)로 구성된다.
- <24> 음성/목음분류부(210)는 고정 코드북 이득(Fixed Code-Book Gain : FCBG), 적응 코드북 이득(Adaptive Code-Book Gain : ACBG), 잡음대 신호비(Noise to Signal Rate : NSR), 및 피치 지연(Pitch Delay)를 이용하여 부호화된 비트열의 입력 파라미터로부터 입력 프레임을 음성과 목음으로 분류한다. 이 때, 음성/목음분류부(210)는 입력 비트열에 대한 고정 코드북 이득값과 적응 코드북 이득값이 설정되어 있는 제1문턱값보다 크고, 잡음대신호비와 피치지연값이 설정되어 있는 제2문턱값보다 작으면 입력되는 비트열에 대응하는 프레임을 음성으로 분류한다.

- <25> G.729A의 피치지연값은 음성이 아닌 구간에서는 그 변화가 크게 나타난다. 이러한 성질을 이용하면 음성인 구간과 음성이 아닌 구간을 분류해낼 수 있다. 도 3에는 일정 구간 동안(두 프레임) 입력으로 들어오는 G.729A 신호의 피치지연값의 최대값과 최소값의 차이 및 사용한 음성을 도시한 도면이다. 도 3을 참조하면, 음성이 존재하는 구간에서는 G.729A 신호의 피치지연값의 최대값과 최소값의 차이가 상당히 작지만 음성이 존재하지 않는 구간에서는 그 차이값이 무척 크게 나타난다. 음성/묵음분류부(210)는 이와 같은 피치지연값의 성질을 이용해서 음성구간과 묵음구간을 분리한다.
- <26> 또한, 적응 코드북 이득값은 그 변화가 심하지만, 한 프레임 내에서 최소값만을 이용할 경우, 음성구간과 묵음구간을 분류할 수 있다. 도 4는 적응 코드북 이득값의 프레임별 최소값을 도시한 도면이다. 도 4를 참조하면, 적응 코드북 이득값의 프레임별 최소값은 음성이 있는 구간에서는 큰 값을 가지고 음성이 없는 구간에서는 작은 값을 가진다. 따라서, 음성/묵음분류부(210)는 적응 코드북 이득값의 프레임별 최소값에 대해 설정된 문턱값을 기초로 음성구간과 묵음구간을 구분할 수 있다.
- <27> 한편, 일반적으로 음성부호화기에서는 고정 코드북 이득값이 음성의 모양과 가장 유사한 형태를 가지게 된다. 이러한 고정 코드북 이득값에 의해 음성을 음성구간과 묵음구간으로 분류하는 것이 가능하다. 즉, 고정 코드북 이득값에 대해 소정의 문턱값을 설정하여 설정된 문턱값을 기준으로 음성과 묵음을 분류한다. 그러나, 입력되는 G.729A의 비트열을 생성한 음성입력에 잡음이 존재할 경우에는 고정 코드북 이득값을 이용한 음성과 묵음의 분류는 양호지 않은 결과를 낳게 된다. 도 5는 단일음성신호에 대해서 깨끗한 신호와 백색잡음을 섞은 신호에 대해서 G.729A의 고정 코드북 이득값을 도시한 도면이다. 도 5를 참조하면, 하단의 깨끗한 신호가 백색잡음이 섞이지 않는 신호의 고정 코드북 이득값이고, 상단의 신호가 백색잡음을 섞은 신호의

고정 코드북 이득값이다. 도 5로부터 백색잡음이 섞였을 경우에 잡음의 크기 때문에 음성구간과 묵음구간을 나누는 기준을 설정하기 곤란하다는 것을 확인할 수 있다. 이와 같이 잡음이 섞였을 경우에 고정 코드북 이득값을 이용해 음성을 분류하는 것은 바람직하지 않다. 따라서, 고정 코드북 이득값은 잡음대신호비(noise-to-signal ratio, NSR) 수치가 아주 낮을 경우에만, 즉 노이즈가 거의 섞이지 않았다고 판단되는 음성 프레임에 대해서만 음성 구간과 묵음 구간을 분리하기 위해 사용된다. 그리고, NSR이 아주 높은 경우에는 잡음이 많이 섞였기 때문에 묵음 구간으로 결정한다.

<28> 유/무성음분류부(220)는 적응 코드북 이득을 이용하여 음성으로 인식된 입력 파라미터로부터 입력 프레임을 유성음과 무성음으로 분류한다. 유/무성음분류부(220)는 음성/묵음분류부(210)에 의해 음성으로 분류된 입력 비트열에 대한 적응 코드북 이득값이 설정되어 있는 문턱값보다 크면 입력 비트열에 대응하는 프레임을 변화음 또는 유성음으로 분류하고, 문턱값보다 작으면 무성음으로 분류한다. 즉, 유/무성음분류부(220)는 도 4를 참조하여 설명한 적응 코드북 이득값의 프레임별 최소값에 대해 음성과 묵음구간을 구분하기 위한 문턱값보다 조금 더 큰 문턱값에 의해 유성음 구간과 무성음 구간을 구분한다. 이 때, 문턱값은 여러 가지 음성에 대해서 적용 가능하고, 잡음이 섞였을 때도 음성을 잘 분류할 수 있도록 설정되어야 한다.

<29> 유성음/변화음분류부(230)는 과거 프레임의 클래스를 기초로 유/무성음분류부(220)에 의해 변화음 또는 유성음으로 분류된 입력 파라미터로부터 입력 프레임을 유성음과 변화음으로 분류한다. 유성음/변화음분류부(230)는 변화음 또는 유성음으로 인식된 입력 비트열에 대한 과거 프레임의 클래스와 현재 프레임의 클래스가 동일하면 유성으로 분류하고, 상이하면 변화음으로 분류한다.

<30> 유성음분류부(240)는 적응 코드북 이득과 피치지연을 이용하여 유성음/변화음분류부(230)에서 유성음으로 분류된 입력 파라미터로부터 입력 프레임을 일정한 유성음과 일정하지 않은 유성음으로 분류한다. 적응 코드북 이득을 이용하는 경우에, 유성음분류부(240)는 프레임 내의 전체 적응 코드북 이득값이 일정한가를 파악하여 유성음을 일정하지 않은 유성음과 일정한 유성음으로 분류할 수 있다. 피치지연을 이용하는 경우에, 유성음분류부(240)는 피치지연값의 최대값과 최소값의 차이가 아주 작은 경우는 일정한 피치 지연값을 갖는 경우이므로 이에 의해 일정한 유성음과 일정하지 않은 유성음을 분류한다.

<31> 전송률결정부(250)는 각각의 분류부(210 내지 240)에서 분류된 프레임에 대해 전송률과 형태를 결정한다. 이 때, 전송률결정부(250)는 표 2에 기재된 모드에 따라 프레임에 대한 전송률과 형태를 결정한다. 전송률결정부(250)는 모드 1, 2, 3에 대해서는 프레임을 분류할 때 각기 다른 문턱값을 적용한다. 본 발명에서는 분류의 간단화를 위해 잡음과 무성음을 구분하지 않는다.

<32> 【표 2】

모드	프레임 클래스	레이트 $\frac{1}{8}$	레이트 $\frac{1}{4}$	레이트 $\frac{1}{2}$	레이트 1
0	목음, 잡음	√		√	√
	무성음			√	√
	변화음			√	√
	일정하지 않은 유성음				√
	일정한 유성음				√
1, 2, 3	목음, 잡음	√	√	√	
	무성음		√	√	
	변화음		√	√	√
	일정하지 않은 유성음			√	√
	일정한 유성음			√	√

<33> 도 6은 본 발명에 따른 음성의 상호부호화시 전송률 결정 방법의 수행과정을 도시한 흐름도이다.

<34> 도 6을 참조하면, 음성/묵음분류부(210)는 부호화된 비트열의 입력 파라미터로부터 고정 코드북 이득, 적응 코드북 이득, 잡음대 신호비, 및 피치지연 중 적어도 하나를 이용하여 입력 프레임의 음성과 묵음으로 분류한다(S600). 유/무성음분류부(220)는 음성으로 인식된 입력 파라미터로부터 적응 코드북 이득을 이용하여 입력 프레임을 변화음/유성음과 무성음으로 분류한다(S610). 유성음/변화음분류부(230)는 변화음/유성음으로 인식된 입력 파라미터로부터 과거 프레임 클래스를 이용하여 입력 프레임을 유성음과 변화음으로 분류한다(S620). 유성음분류부(240)는 유성음으로 인식된 입력 파라미터로부터 적응 코드북 이득 또는 피치 지연을 이용하여 입력 프레임을 일정하지 않는 유성음과 일정한 유성음으로 분류한다(S630). 전송률결정부(250)는 분류된 프레임이 속하는 클래스에 대해 설정되어 있는 전송률과 형태를 기초로 입력 프레임의 전송률을 결정한다(S640).

<35> 본 발명은 또한 컴퓨터로 읽을 수 있는 기록매체에 컴퓨터가 읽을 수 있는 코드로서 구현하는 것이 가능하다. 컴퓨터가 읽을 수 있는 기록매체는 컴퓨터 장치에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록장치를 포함한다. 컴퓨터가 읽을 수 있는 기록매체의 예로는 ROM, RAM, CD-ROM, 자기 테이프, 플로피디스크, 광데이터 저장장치 등이 있으며, 또한 캐리어 웨이브(예를 들어 인터넷을 통한 전송)의 형태로 구현되는 것도 포함한다. 또한 컴퓨터가 읽을 수 있는 기록매체는 네트워크로 연결된 컴퓨터 장치에 분산되어 분산방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수 있다.

<36> 이상에서 본 발명의 바람직한 실시예에 대해 도시하고 설명하였으나, 본 발명은 상술한 특정의 바람직한 실시예에 한정되지 아니하며, 청구범위에서 청구하는 본 발명의 요지를 벗어남이 없이 당해 발명이 속하는 기술분야에서 통상의 지식을 가진 자라면 누구든지 다양한 변형 실시가 가능한 것은 물론이고, 그와 같은 변경은 청구범위 기재의 범위 내에 있게 된다.

【발명의 효과】

<37> 본 발명에 따른 음성의 상호부호화시 전송률 결정 장치 및 방법에 의하면, CELP 기반의 코덱으로 부호화된 신호에서 SMV신호로 상호부호화시 입력 비트열의 파라미터를 이용함으로써, 프레임을 용이하게 분류하고 전송률 결정 과정을 간단하게 구현할 수 있으며, 전체적인 계산량이 감소될 수 있다.

【특허청구범위】

【청구항 1】

부호화된 비트열의 입력 파라미터에 대응하는 고정 코드북 이득, 적응 코드북 이득, 잡음대 신호비, 및 피치지연 중에서 적어도 하나의 값에 대해 설정된 제1문턱값을 기초로 상기 입력 파라미터로부터 입력 프레임을 음성과 묵음으로 분류하는 음성/묵음분류부;

상기 적응 코드북 이득에 대해 설정된 제2문턱값을 기초로 음성으로 분류된 상기 입력 파라미터로부터 상기 입력 프레임을 유성음과 무성음으로 분류하는 유/무성음분류부;

과거 프레임의 클래스를 기초로 상기 유/무성음분류부에 의해 유성음으로 분류된 상기 입력 파라미터로부터 상기 입력 프레임을 유성음과 변화음으로 분류하는 유성음/변화음분류부;

상기 적응 코드북 이득의 변화량 또는 상기 피치지연의 최대값과 최소값의 차이에 대해 설정된 제3문턱값을 기초로 상기 유성음/변화음분류부에 의해 유성음으로 분류된 상기 입력 파라미터로부터 상기 입력 프레임을 일정한 유성음과 일정하지 않은 유성음으로 분류하는 유성음 분류부; 및

상기 분류결과에 대응하는 클래스에 대해 설정되어 있는 전송률과 형태를 기초로 상기 입력 프레임에 대한 전송률을 결정하는 전송률결정부;를 포함하는 것을 특징으로 하는 음성 상호부호화시 전송률 결정 장치.

【청구항 2】

(a) 부호화된 비트열의 입력 파라미터로부터 소정의 제1문턱값을 기초로 입력 프레임을 음성과 묵음으로 분류하는 단계;

(b) 상기 입력 파라미터에 대응하는 적응 코드북 이득에 대해 설정된 제2문턱값을 기초로 음성으로 분류된 상기 입력 파라미터로부터 상기 입력 프레임을 유성음과 무성음으로 분류하는 단계;

(c) 과거 프레임의 클래스를 기초로 상기 유/무성음분류부에 의해 유성음으로 분류된 상기 입력 파라미터로부터 상기 입력 프레임을 유성음과 변화음으로 분류하는 단계;

(d) 상기 적응 코드북 이득의 변화량 또는 상기 입력 프레임에 대한 상기 피치지연값의 최대값과 최소값의 차이값에 대해 설정된 소정의 제3문턱값을 기초로 상기 유성음/변화음분류부에 의해 유성음으로 분류된 상기 입력 파라미터로부터 상기 입력 프레임을 일정한 유성음과 일정하지 않은 유성음으로 분류하는 단계; 및

(e) 상기 분류결과에 대응하는 클래스에 대해 설정되어 있는 전송률과 형태를 기초로 상기 입력 프레임에 대한 전송률을 결정하는 단계;를 포함하는 것을 특징으로 하는 음성 상호부호화시 전송률 결정 방법.

【청구항 3】

제 2항에 있어서,

상기 (a)단계는 상기 입력 파라미터에 대응하는 상기 적응 코드북 이득에 대해 설정된 상기 제1문턱값을 기초로 상기 입력 파라미터로부터 상기 입력 프레임을 음성과 묵음으로 분류하는 것을 특징으로 하는 음성 상호부호화시 전송률 결정 방법.

【청구항 4】

제 3항에 있어서,

상기 제1문턱값은 상기 제2문턱값보다 작게 설정되는 것을 특징으로 하는 음성 상호부호화시 전송률 결정 방법.

【청구항 5】

제 2항에 있어서,

상기 (a)단계는 상기 입력 파라미터의 피치지연값의 최대값과 최소값의 차이값에 대해 설정된 소정의 제4문턱값을 기초로 상기 입력 파라미터로부터 상기 입력 프레임을 음성과 묵음으로 분류하는 것을 특징으로 하는 음성 상호부호화시 전송률 결정 방법.

【청구항 6】

제 5항에 있어서,

상기 제4문턱값은 상기 제3문턱값보다 크게 설정되는 것을 특징으로 하는 음성 상호부호화시 전송률 결정 방법.

【청구항 7】

제 2항에 있어서,

상기 (a)단계는 상기 입력 파라미터에 대응하는 고정 코드북 이득에 대해 설정된 제5문턱값을 기초로 상기 입력 파라미터로부터 입력 프레임을 음성과 묵음으로 분류하는 것을 특징으로 하는 음성 상호부호화시 전송률 결정 방법.

【청구항 8】

제 7항에 있어서,

상기 입력 프레임에 대한 신호대잡음비는 소정의 제6문턱값보다 작은 것을 특징으로 하는 음성 상호부호화시 전송률 결정 방법.

【청구항 9】

(a) 부호화된 비트열의 입력 파라미터로부터 소정의 제1문턱값을 기초로 입력 프레임을 음성과 묵음으로 분류하는 단계;

(b) 상기 입력 파라미터에 대응하는 적응 코드북 이득에 대해 설정된 제2문턱값을 기초로 음성으로 분류된 상기 입력 파라미터로부터 상기 입력 프레임을 유성음과 무성음으로 분류하는 단계;

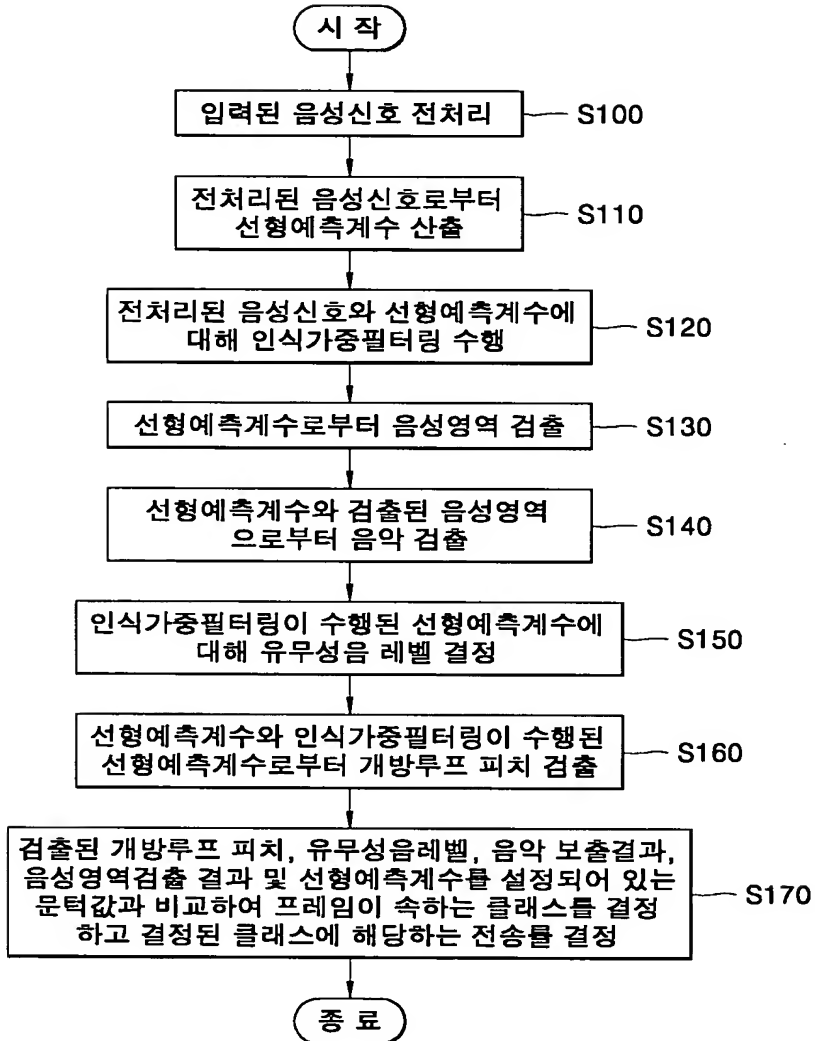
(c) 과거 프레임의 클래스를 기초로 상기 유/무성음분류부에 의해 유성음으로 분류된 상기 입력 파라미터로부터 상기 입력 프레임을 유성음과 변화음으로 분류하는 단계;

(d) 상기 적응 코드북 이득의 변화량 또는 상기 입력 프레임에 대한 상기 피치지연값의 최대값과 최소값의 차이값에 대해 설정된 소정의 제3문턱값을 기초로 상기 유성음/변화음분류부에 의해 유성음으로 분류된 상기 입력 파라미터로부터 상기 입력 프레임을 일정한 유성음과 일정하지 않은 유성음으로 분류하는 단계; 및

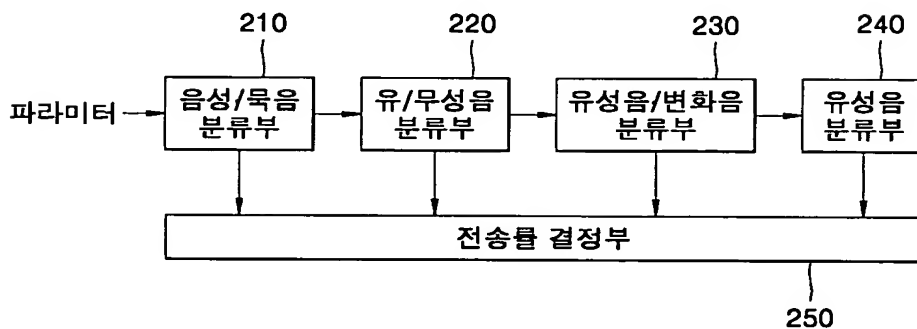
(e) 상기 분류결과에 대응하는 클래스에 대해 설정되어 있는 전송률과 형태를 기초로 상기 입력 프레임에 대한 전송률을 결정하는 단계;를 포함하는 것을 특징으로 하는 음성 상호부호화시 전송률 결정 방법을 컴퓨터에서 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체.

【도면】

【도 1】

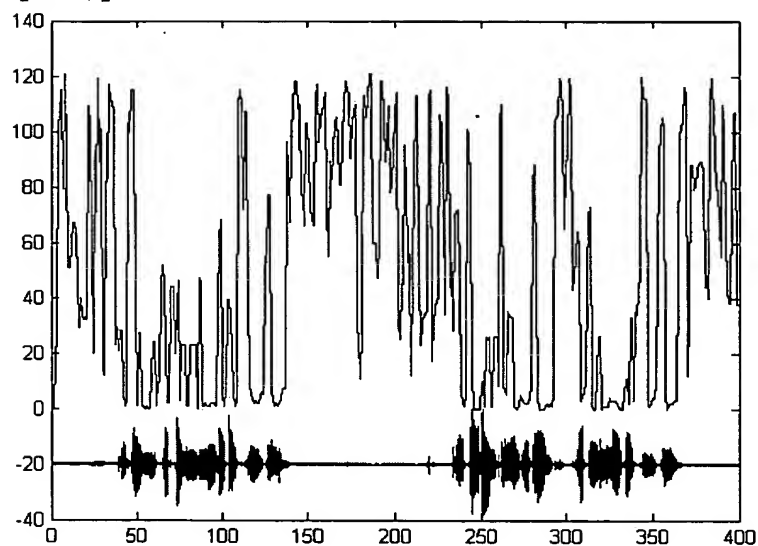


【도 2】

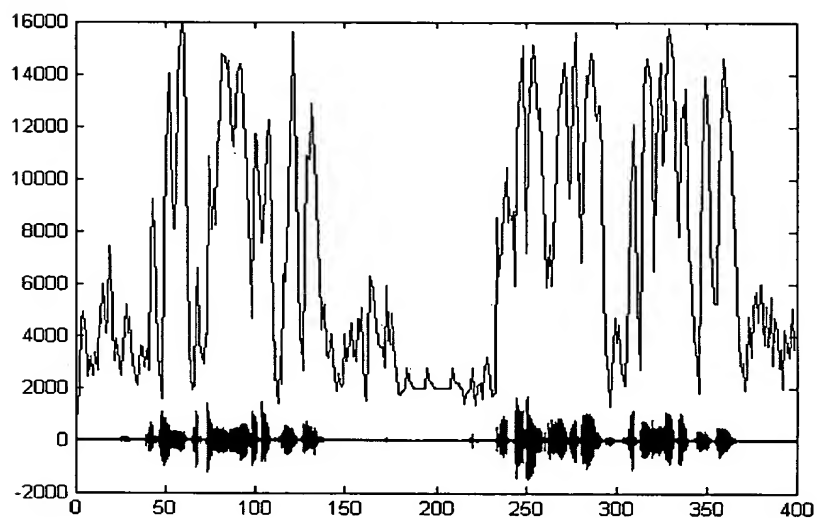


BEST AVAILABLE COPY

【도 3】

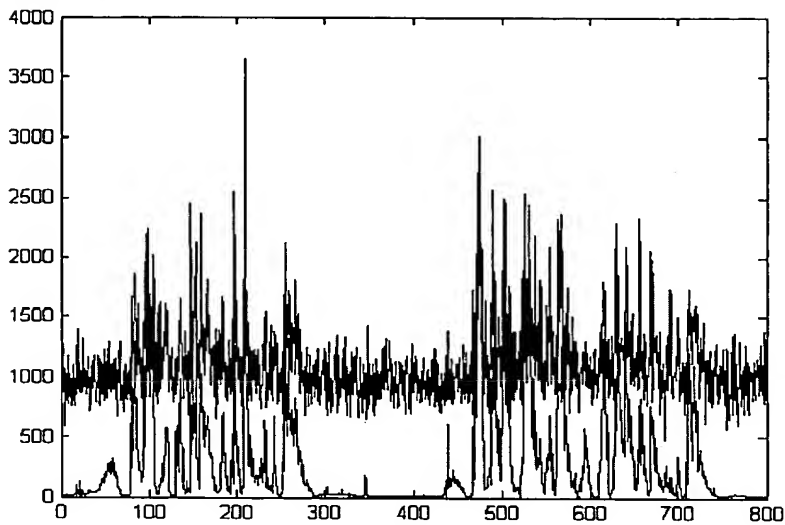


【도 4】

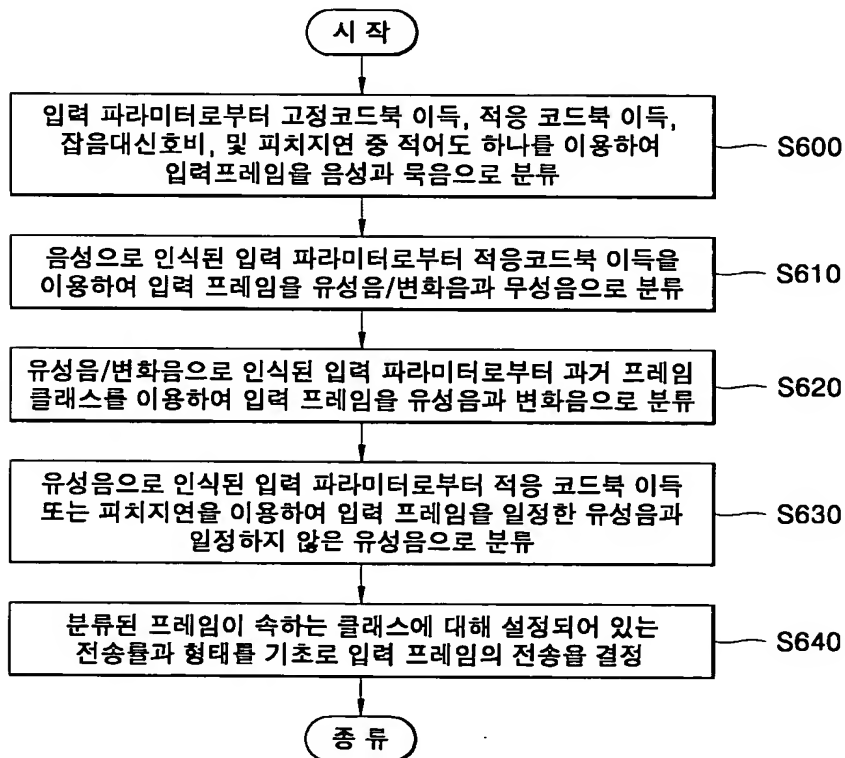


BEST AVAILABLE COPY

【도 5】



【도 6】



BEST AVAILABLE COPY